

Rindermann, Heiner
**Bedingungs- und Effektvariablen in der Lehrevaluationsforschung:
Konzeption und Prüfung des Münchner multifaktoriellen Modells der
Lehrveranstaltungsqualität**

Unterrichtswissenschaft 27 (1999) 4, S. 357-380



Quellenangabe/ Reference:

Rindermann, Heiner: Bedingungs- und Effektvariablen in der Lehrevaluationsforschung: Konzeption und Prüfung des Münchner multifaktoriellen Modells der Lehrveranstaltungsqualität - In: Unterrichtswissenschaft 27 (1999) 4, S. 357-380 - URN: urn:nbn:de:0111-opus-77364 - DOI: 10.25656/01:7736

<https://nbn-resolving.org/urn:nbn:de:0111-opus-77364>

<https://doi.org/10.25656/01:7736>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Digitalisiert

Unterrichtswissenschaft

Zeitschrift für Lernforschung
27. Jahrgang / 1999 / Heft 4

Thema:

Quantitative und qualitative Analysen in der Lehr-Lernforschung: Perspektiven zur Integration

Verantwortlicher Herausgeber:
Alexander Renkl

- Alexander Renkl:
Quantitative und qualitative Analysen in der Lehr-Lernforschung:
Perspektiven zur Integration – Eine Einführung 290
- Philipp Mayring:
Qualitativ orientierte Forschungsmethoden
in der Unterrichtswissenschaft – ein Anwendungsbeispiel
aus der Lernstrategieforschung 292
- Alexander Renkl:
Jenseits von $p < .05$: Ein Plädoyer für Qualitatives 310
- Detlev Leutner:
Hypothesenprüfung versus interpretative Exploration:
Die endlose Debatte zur Funktion quantitativer und qualitativer
Analysen in der Lehr-Lernforschung (Kommentar) 323
- Kai Schnabel:
Wissenschaft ist Abstraktion, aber Abstraktes
ist zu nix zu gebrauchen (Kommentar) 333

Allgemeiner Teil

- Hans-Joachim Vogler:
Zur Politisierung des Blicks im Staatsbürgerkundeunterricht der DDR 337
- Heiner Rindermann:
Bedienungs- und Effektvariablen in der Lehrevaluationsforschung:
Konzeption und Prüfung des Münchner multifaktoriellen Modells
der Lehrveranstaltungsqualität 357

- Buchbesprechungen** 381
- 289

Heiner Rindermann

Bedingungs- und Effektivvariablen in der Lehrevaluationsforschung: Konzeption und Prüfung des Münchner multifaktoriellen Modells der Lehrveranstaltungsqualität¹

Concept and Examination of the
Munich Multifactorial Model of Course Quality

Studentische Hochschulgruppen, Fachbereiche und universitäre Institutionen beschäftigen sich seit Beginn der 90er Jahre mit Lehrevaluationen. Eine theoretische Basis ist häufig aber nicht erkennbar. Deshalb soll hier ein multifaktorielles hochschuldidaktisches Konzept vorgestellt und geprüft werden, welches universitäre Lehre als Ergebnis von Bedingungsvariablen auf seiten des Dozenten, der Studierenden und der Rahmenbedingungen auffaßt. Das Handeln des Dozenten wird in didaktische und soziale Komponenten unterschieden, studentisches Handeln in Veranstaltungsbe teiligung und Arbeitshaltung, Rahmenbedingungen in Veranstaltungstyp und -the men. Als Produktvariablen der Lehreffektivität lassen sich u.a. Lernerfolg und die all gemeine Kursqualität erheben. Dieses Konzept wird auf seine theoretische Stringenz und empirische Bewährung mittels Regressionen und Strukturgleichungsmodellen überprüft.

Student groups, academic organizations and faculties have dealt with the problem of student evaluation of teaching since the early nineties. A theoretical base has, so far, not been established in many projects. Therefore, a multi-factored concept of university didactics will be presented and examined which regards university teaching as a result of conditional variables found in instructors, students and external conditions. For example, instructor-variables can be classified according to both didactic and social components. The way students participate in courses can be separated from their approach to work. External conditions can be divided according to course type and course topic. In addition learning success and a general quality of a course should be charted as product variables of teaching effectiveness. This concept will be checked for its theoretical clearness and its empirical worth by using stepwise regressions and structural equation models.

¹ Herrn Dipl.-Psych. Jürgen Kohler möchte ich für die Durchführung einer Evaluationsstudie an der Schule für Logopädie in Reichenau danken, ebenso Herrn PD Dr. Norbert Tanzer für seine hilfreiche Unterstützung bei LISREL-Analysen.

1. Bisherige Ansätze zur Generierung von Skalen zur Veranstaltungskritik

Mit steigendem öffentlichen Interesse an der Qualität der Lehre sind in den letzten Jahren an deutschsprachigen Universitäten Instrumente zur Beurteilung der Lehre entwickelt worden (s. Dokumentation von Reissert, 1992, und Hage, 1996). Nur selten stellen die Instrumente Fortentwicklungen älterer überprüfter Verfahren dar, in der Regel orientierten sich neue Verfahren mehr an Anforderungen der Praxis und weniger an theoretischen Modellen.

Unter den Itemgenerierungsverfahren für Fragebögen lassen sich prinzipiell folgende Methoden unterscheiden:

1. *Befragung*: Offene Befragung von Studierenden, Dozenten oder Experten (Hochschuldidaktiker) nach relevanten Kriterien guter Lehre. Aus den eruierten allgemeinen oder fach- und veranstaltungsspezifischen Kriterien lassen sich Items formulieren und ein Fragebogen zusammenstellen.
2. *Synkretistisch*: Bislang vorhandene Verfahren und Ergebnisse der Lehrevaluationsforschung dienen der Auswahl und Kombination geeigneter Items und Skalen.
3. *Gremienbeschluß*: Hochschulgremien handeln intern einen Fragebogen und dessen Items aus.
4. *Theoretisch*: Die Itemgenerierung basiert auf schon existierenden oder für die Fragebogenkonstruktion entwickelten Theorien und Lehrzielkonzeptionen.
5. *Autor*: In Eigenregie entwickelt eine Person ein Instrument anhand impliziter oder expliziter Annahmen.

Jedes der skizzierten Verfahren zeichnet sich durch spezifische Vor- und Nachteile aus. Während im *Befragungs-Modell* die Generierungskriterien explizit sind, bleiben die letztendlichen Auswahlverfahren durch einen oder mehrere Autoren oft unklar, weil bei Befragungen zu viele und nicht „itemfertige“ Nennungen anfallen, die noch zu Items destilliert werden müssen. Hierzu müßte wie im *synkretistischen Modell* ein strukturierendes Konzept herangezogen werden – und sei es nur eine Kombination nach Ähnlichkeit und anschließende Auszählung –, um aus der Fülle der Anregungen die relevantesten Aspekte herausfinden zu können. Vorteil des ersten Verfahrens ist die Nähe zur Hochschulpraxis und zu den Instrument-Nutzern, beim zweiten die Wahrscheinlichkeit, auf empirisch bewährte wie auch inhaltlich bedeutsame Items zu stoßen. Das synkretistische Verfahren zeichnet sich durch geringen Aufwand aus, allerdings sind die Fragebögen nicht immer direkt übertragbar (US-amerikanische Bögen auf deutsche Verhältnisse, Post-68er-Entwicklungen² auf die heutige Hochschulsituation, andere Studienfächer).

² Beispiel: Müller-Wolf (1977) mit Skalen „demokratisches Lehrverhalten (vs. autoritäres)“ oder „ideologiekritisch-progressive wissenschaftstheoretische Haltung“. (Vgl. a. Leitner, 1998).

In *Gremien* entwickelte Instrumente stoßen in der Praxis auf große Akzeptanz, wenn sich alle relevanten Entscheidungs- und Nutzergruppen im fertigen Inventar wiederfinden können. Allerdings kann das Ergebnis eines solchen Verfahrens ein unstrukturierter, weder an theoretischen noch empirischen Kriterien orientierter Kompromißbogen sein (vgl. Marsh & Roche, 1997). Im Extremfall sind die Ergebnisse unverwendbar, weil Items, die ein kritisches Ergebnis zu Tage fördern mögen, im Aushandlungsverfahren eliminiert wurden – übrig bleiben dann nur äußerliche Aspekte des Kurses wie Anzahl der Ausfälle und Länge der Referate. Anhand *theoretischer Modelle* erstellte Instrumente halten den Vorteil eines in sich geschlossenen Konzepts bereit, darüber hinaus verfügen sie über einen höheren „wissenschaftlichen Verteidigungswert“. Aus grundlegenden theoretischen Annahmen oder Lehrzielkonzeptionen (z.B. Schweer, 1998; Leitner, 1998; Krempkow, 1998) werden einzelne, in einem Fragebogen einsetzbare Items abgeleitet, die inhaltlich durch die Theorien begründet werden. Der theorienbasierte Fragebogen eignet sich jedoch nur dann für den Einsatz, falls das Konzept dem Hochschulunterricht gerecht wird, für den das Instrument entwickelt wurde. Für das *Autorenverfahren* sprechen Unkompliziertheit und Freiheitsgrade, dagegen die Intransparenz der Kriterien und eine fehlende Orientierung an theoretischen Modellen oder empirischen Befragungen.

2. Theoretische Ansätze in der Lehrevaluationsforschung

Die meisten sich an Theorien orientierenden Generierungsverfahren bedienen sich mehrerer unabhängiger theoretischer Grundannahmen und kombinieren diese. Diese enthalten in der Regel verdeckt normative Bestandteile. Beispielsweise ist der Hypothese „Wenn der Dozent klar spricht, dann verstehen und lernen Studierende mehr“ indirekt entnehmbar, daß „verstehen“ und „lernen“ und damit auch „klar sprechen“ als wünschenswerte Aspekte des Hochschulunterrichts gelten. Solche kryptonormativen Gehalte (Gloebe, 1986, S. 417) sind nicht nur theoretischen Grundlagen, sondern auch Fragebögen und deren Items entnehmbar. Die in Lehrinventaren behandelten Aspekte formulieren implizit Annahmen, wie der Hochschulunterricht aussehen *sollte* (z.B. gut strukturiert, motivierender Dozent; ausf. Wilson, 1988). Da die Items Zielvorstellungen und Anforderungen an die Lehre widerspiegeln, können sie durch ihre Verwendung in Inventaren auf die Lehre rückwirken, während nicht thematisierte Aspekte möglicherweise ungerechtfertigt marginalisiert werden.

Drei Beispiele sollen hier kritisch besprochen werden: Das von Westermann et al. (1993, 1996) und von Kramis (1990) sowie das Modell von Helmke (1996), das bislang aber noch nicht als Grundlage für Lehrinventare diente.

2.1 Der Ansatz von Westermann et alii

Die Forscher(innen)gruppe um Westermann (et al. 1993; et al. 1996; Spies, Westermann, Heise & Schiffler, 1996) orientierte sich bei der Konstruktion

eines Fragebogens zur Erhebung der Zufriedenheit mit der Lehre an verschiedenen Theorien:

- *Konzept der Arbeitszufriedenheit*: Die Einstellung einer Person zu ihrer Arbeit ist Folge einer Linearkombination der Valenz der Objektattribute und der Stärke der subjektiven Überzeugungen, daß die Objekte diese Attribute besitzen.
- *Person-Environment-Fit-Theorie*: Die Zufriedenheit einer Person ist abhängig von der subjektiven Passung zwischen ihren Bedürfnissen, Einstellungen, Fähigkeiten und den Merkmalen der studienbezogenen Umwelt.
- Zusätzlich werden für Teilaspekte humanistische Bedürfnishierarchien, Lernzieltaxonomien sowie die Job-Characteristics-Theorie und Handlungsphasentheorie herangezogen.

Die Fragebogenentwicklung richtete sich an den Konzepten der (Arbeits-/ Studien-) Zufriedenheit und der Person-Umwelt-Passung aus. Die Zufriedenheit gilt als interaktives Produkt von Personen- und Umweltmerkmalen. Zwar wurden die theoretischen Grundlagen der Skalenkonstruktion für Lehr- und Lehrveranstaltungsinventare dargelegt. Allerdings bleibt unklar, inwieweit diese Theorien für den tatsächlichen Skalengenerierungsprozeß ausschlaggebend waren. Die späteren Items und Dimensionen des Lehrveranstaltungsinventars lassen sich zumindest nicht den einzelnen theoretischen Annahmen zuordnen.

Nicht die Qualität der Lehre stellt bei Westermann et alii das zu analysierende Merkmal dar, sondern die Zufriedenheit der Studierenden mit der Lehre. Didaktische Theorien fehlen, die Hypothesen über einzelne Lehrformen und Unterrichtstechniken in Zusammenhang mit verschiedenen Lehrerfolgskriterien formulieren. Tatsächlich scheint aber der Mehrzahl der Items eine implizite didaktische Theorie mit dem zumindest partiell vorhandenen Anspruch zugrundezuliegen, das Lehrgeschehen und nicht nur die Zufriedenheit der Studierenden zu beschreiben. Das Zufriedenheits-Paradigma bedarf theoretisch wie auch praktisch einer Ergänzung durch instruktionale Theorien, um einen fruchtbaren Beitrag zur Beschreibung, Erklärung und Verbesserung universitärer Lehre zu bieten. Als ein *Kriterium des Lehrerfolges* (neben anderen wie dem Lerngewinn) ist die studentische Veranstaltungszufriedenheit jedoch eine bedeutsame Größe.

2.2 Gütekriterien von Kramis

Kramis Konzeption (1990) zeichnet sich aus durch die Orientierung an drei postulierten „grundlegenden Gütekriterien für Unterricht“³ (1990, S. 279), die er in Anlehnung an theoretische didaktische Modelle sowie instruktionspsychologische Metaanalysen konstruiert hat. Diese sind:

³ Kramis spricht von „Unterricht“ im tertiären und sekundären Bildungssektor. Im vorliegenden Beitrag wird „Lehre“ und „Unterricht“ synonym gebraucht (vgl. a. den einheitlichen englischen Begriff „instruction“).

- Bedeutsamkeit der gewählten Unterrichtsinhalte und Ziele,
- Effizienz der gewählten Lernorganisation, Lernaktivitäten und Medien und
- gutes Lernklima.

Den genannten fundamentalen und notwendigen Gütekriterien ordnet Kramis 36 didaktische Prinzipien zu und leitet von diesen wiederum Items ab. Faktorenanalytisch bestätigen sich die drei Unterrichts-Gütekriterien, die Skalen sind homogen (Trennschärfen und Cronbach- α). Die Analysen basieren nur auf einer Stichprobe von N=110 Studierenden (7 Veranstaltungen).

Bei Kramis sind einzelne Items den didaktischen Theorien genau zuzuordnen. Die empirische Abstützung ist jedoch sowohl auf seiten der Stichprobe (relativ klein und nur Lehramtsstudium) als auch auf seiten der verwendeten statistischen Verfahren (Stabilitäts- und Validitätsanalysen stehen aus) schwach. Seine Konzeption stellt eher ein präskriptives Modell („Leitbilder“) für die Lehre und Lehrende dar, als ein empirisch beschreibendes und Kausalbeziehungen aufzeigendes Bedingungsmodell guten Unterrichts (vgl. Mayr, 1997, S. 244ff.). Unterschieden werden zwar Inhalte, Lehrstrategien (die die Effizienz sichern sollen) und soziale Komponenten, allerdings wird nicht präzise zwischen Rahmenbedingungen (Skala „Bedeutsamkeit“ umfaßt sowohl vorgegebene Themen als auch Unterrichtsformen) und dem Dozenten- oder Studentenhandeln differenziert. Das Vorinteresse am Thema, die studentische Arbeitshaltung, die Höhe der Anforderungen und der selbsteingeschätzte Lerngewinn bleiben unthematisiert. Exemplarisch bleibt bei Kramis sicherlich die enge Verzahnung von Theorie, Fragebogenkonstruktion und Lehrpraxis.

2.3 Angebots-Nutzungs-Modell des Studienerfolgs von Helmke (1996)

Helmke (1996) stellt ein allgemeines Modell der Studienbedingungen und des Studienerfolgs vor (s.a. Helmke & Schrader, 1998). Er unterscheidet hierbei Lehrkontext, Lehrangebot, Lehrperson, studentische Determinanten und individuelle Lernvoraussetzungen als Bedingungen der nominalen und aktiven Nutzung des Lehrangebots. Auf der Effektseite wird zwischen kognitiven und nichtkognitiven Wirkungen und langfristigen Effekten wie Berufserfolg und Lebensbewältigung differenziert. Diese stehen in einem Wechselwirkungsprozeß zur Perzeption und Evaluation.

Bislang wurde dieses Modell noch nicht als Grundlage für Lehrevaluationen oder Skalenkonstruktionen herangezogen. Für ein Evaluationsmodell der universitären Lehre wäre das Angebots-Nutzungs-Modell zu allgemein, hier haben der Wissenschaftsrat (1996) und Webler (1996) detailliertere Programme vorgelegt. Es könnte aber als theoretische Grundlage solcher Programme dienen. Modelle, die viele mögliche Bedingungsfaktoren ohne Gewichtung einzelner Pfade in Beziehung setzen, können heuristisch zur Formulierung einzelner Hypothesen genutzt werden. Widerlegbar oder empirisch prüfbar

sind sie aber nicht. Die Auswahl der Faktoren und Pfade bei Helmke ist fragwürdig: Warum wirkt in dem veranstaltungsübergreifenden Modell die Nutzung des Lehrangebots nicht auf die individuellen Lernvoraussetzungen zurück? Wodurch wird begründet, daß die Attraktivität einer Lehrperson als Bedingungsvariable aufgeführt wird? Warum fehlen Studienabbruch und -dauer als Studienerfolgskriterien?

3. Das Münchner multifaktorielle Modell der Lehrveranstaltungsqualität

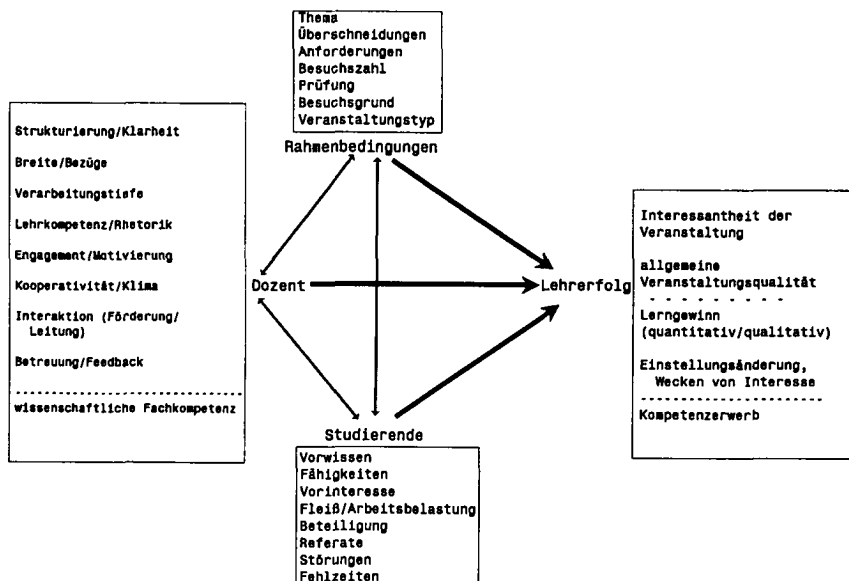
3.1 Entwicklung und Struktur des Modells

Das Münchner multifaktorielle Modell der Lehrveranstaltungsqualität basiert auf verschiedenen empirischen Studien und unterschiedlichen theoretischen Annahmen (vgl. Rindermann & Amelang, 1994; Rindermann, 1996, 1997a): Zu diesen zählen Befragungen von Studierenden und Lehrenden, Auswertungen bisheriger Verfahren, Orientierung an Lehrzielen sowie Berücksichtigung von unterrichtsdidaktischen Theorien. Zur empirischen Überprüfung ließen sich Resultate statistischer Verfahren (Faktoren-, Reliabilitäts- und Validitätsanalysen etc.) nutzen, die an Stichproben unterschiedlicher Universitäten und Studiengänge gewonnen wurden.

Ein synthetisches Konstruktionsverfahren berücksichtigt verschiedene Quellen bei der Itementwicklung; zur Generierung und anschließenden Überprüfung wird ein theoretisch-empirisches Vorgehen gewählt. Ein multifaktorielles Modell der Lehrveranstaltungsqualität bildet den konzeptionellen Rahmen des Entwicklungsverfahrens, ein Modell, welches die Qualität der Lehre (und nicht die Zufriedenheit!) als Ergebnis des Zusammenspiels mehrerer, sich gegenseitig beeinflussender Faktoren ansieht: des Handelns des Dozenten, des Handelns der Studierenden und der Wirkungen der Rahmenbedingungen. Die Vernachlässigung einer dieser drei Komponenten würde weder der Beschreibung des Lehrgeschehens noch der Erklärung des Lehrerfolges gerecht. Lehre stellt eine intentionale interpersonale Aktivität mit dem Ziel einer Modifikation von kognitiven und nichtkognitiven Strukturen dar. Zum Gelingen dieses interaktiven Prozesses müssen die Dozenten und Studenten entsprechende Voraussetzungen erfüllen. Dieser Prozeß ist in gesellschaftliche, curriculare und institutionelle Rahmenbedingungen eingebettet.

Im Modell werden die Komponenten „Dozent“, „Studierende“, „Rahmenbedingungen“ und „Lehrerfolg“ unterschieden (s. Abbildung 1). Erst das Zusammenspiel eines fachlich, didaktisch und kommunikativ kompetenten Dozenten, vorgebildeter, interessierter und engagierter Studenten sowie günstiger Rahmenbedingungen in Form von ausreichend Zeit, Platz, adäquater Anforderungshöhe und inhaltsangemessenen Veranstaltungsformen kann Lehrerfolg mit hoher Wahrscheinlichkeit erzielen.

Abbildung 1:
Multifaktorielles Modell der Lehrveranstaltungsqualität



Die drei Hauptfaktoren werden als Bedingungs- und Prozeß-Variablen mit der Produkt-Variable Lehrerfolg in Beziehung gesetzt (vgl. Dunkin & Barnes, 1986). Sie bewirken den Erfolg der Lehre in Form guter und günstig beurteilter Veranstaltungen und in Form intendierter Veränderungen auf seiten der Veranstaltungsteilnehmer. Einzelne Skalen werden zu Faktoren in einem Determinationsgefüge universitärer Lehre gebündelt. Allerdings wirken Bedingungsfaktoren nicht isoliert voneinander auf den Lehrerfolg (vgl. Weinert, Schrader & Helmke, 1989). Beispielsweise kann der Dozent die Studierenden motivieren und dadurch deren Beteiligung erhöhen. Ebenso steigert eine veranstaltungsbezogene Prüfung den Fleiß der Studenten. Die didaktische, fachliche und soziale Kompetenz der Dozenten wirkt auf die Besuchszahl. Umgekehrt beeinträchtigen schlechte Rahmenbedingungen wie große Besuchszahlen die Betreuung, Störungen behindern die Strukturierung und fehlendes Vorwissen oder geringe studentische Kompetenzen lassen durchschnittliche Anforderungen als hohe erscheinen. Die *Kausalbeziehungen* verlaufen nicht nur zwischen den Hauptfaktoren und den Effektvariablen, sondern auch zwischen und innerhalb der einzelnen Faktoren. Beispielsweise kann sich das Vorwissen der Studenten auf das Auftreten von Störungen auswirken (Weinert et al., 1989, S. 907).

Der Lehrerfolg wird als zeitlich nachgeschaltete Produktvariable aufgefaßt. Allerdings wirken die Erwartungen der Studierenden über den möglichen Lehrerfolg Ende des Semesters und die während der Vorlesungszeit vorgenommenen Bilanzierungen über bisherige Lehrqualität und Lernerfolg auf die Bedingungsvariablen zurück (etwa auf die Besuchszahl).

Im *Gegensatz zu Westermann und Kramis* wird nicht von Interaktivität in Form von Passung zwischen Person- und Umweltmerkmalen oder von gleichzeitiger Erfüllung dreier Unterrichtskriterien als Bedingung von Zufriedenheit oder guten Unterrichts ausgegangen, sondern in Form von wechselseitiger *Beeinflussung* und *Bedingung* dreier Hauptfaktoren guter Lehre und Lehreffektivität. Neben unidirektionalen Effekten sind ebenso auch kompensatorische und wechselseitige, sich gegenseitig fördernde oder hemmende Beeinflussungen zu konstatieren. So hat gleiches Dozentenverhalten vor verschiedenen Studentengruppen oder bei unterschiedlichen Veranstaltungsformen andere Auswirkungen (aptitude-treatment-interaction). Lehrkompetenz zeichnet sich deshalb auch durch adaptives Einsetzen didaktischer Methoden aus (Klauer, 1985; McKeachie, 1997).

Im *Vergleich zu dem von Helmke* (1996) vorgestellten Angebots-Nutzungs-Modell des Studienerfolgs fokussiert das Münchner Modell die Qualität und Effektivität der Lehre in Form von Lehrveranstaltungen. Allgemeine Rahmenbedingungen wie gesellschaftliche und materielle Faktoren (Zulassungssystem, Finanzierung, Alter der Studierenden, Vorbildung) erwiesen sich in bisherigen Analysen der Lehrevaluationsforschung als nicht prognosekräftig für individuelle oder aggregierte Lehrevaluationsresultate (Rindermann, 1997a). Diese sollten nur in ein allgemeines Modell der Studienbedingungen und des Studienerfolges wie von Helmke aufgenommen werden. In Analogie zur „Veranstaltungszufriedenheit“ im Ansatz von Westermann et al. (1993, 1996) wird ein generelles Erfolgskriterium in der Form der „Veranstaltungsqualität“ implementiert. Als Bedingungsvariablen dienen jedoch wie bei Kramis (1990) Dimensionen, die das Lehrgeschehen thematisieren. Bei den Dozentenskalen werden hierzu didaktische Skalen verwendet.

Das Münchner Modell stützt sich vor allem auf Erfahrungen mit dem Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE, Rindermann & Amelang, 1994) und stellt eine Weiterentwicklung dieser Konzeption dar (Rindermann, 1997a). Prädiktor-Kriteriums-Beziehungen werden in ein Konzept der Unterrichtsqualität integriert. Modifizierte Formulierungen unterscheiden nun stärker zwischen den drei Hauptfaktoren Dozent-Studenten-Rahmenbedingungen (so Themarelevanz vs. Vorinteresse vs. Interessantheit der Kursgestaltung), auch zählen die Skalen Interaktion, Betreuung/Feedback, Vorwissen, Redundanz, Störungen und Fehlzeiten nun zum Grundkanon des revidierten Instruments. Zwischen Lernen in quantitativer und qualitativer Form wird unterschieden. Interessenförderung wurde durch ein weiteres Item ergänzt.

Dem Ziel, die wesentlichen Bedingungen und Effekte der Lehre in verschiedenen Studienrichtungen und Veranstaltungsformen zu formulieren, folgt nun die Beschreibung der Skalen und die Begründung ihrer Auswahl durch Theorien und bisherige Resultate der Evaluations- und Lehr-Lern-Forschung (Itemliste des HILVE-II s. Anhang). Anschließend werden Prädiktor-Kriteriums-Beziehungen durch schrittweise Regressionen und Pfadanalysen überprüft.

3.2 Bedingungsfaktor Dozent (Handlungen, Kompetenzen und Einstellungen)

Strukturierung/Klarheit

Ein für die Teilnehmer nachvollziehbarer Aufbau zählt zu den zentralen Merkmalen guten Unterrichts (Bolton et al., 1976; Feldman, 1976; Weinert et al., 1989). Eine klare Struktur fördert den Verständnisprozeß, erleichtert die kognitive Strukturierung der zu lernenden Inhalte und deren späteren Abruf. Die Skala Struktur korreliert in HILVE-Studien mit allen Lehrerfolgsskalen zu $r \geq .50$. Land (1981) sowie Cruickshank und Kennedy (1986) wiesen einen positiven Zusammenhang zwischen klarer Strukturierung und objektivierter studentischer Leistung in der Höhe von $r = .40-.60$ nach (vgl. a. Weinert, 1996). Aus Studenten- und Dozentensicht in verschiedenen Hochschul-Systemen und Unterrichtstraditionen stellt die Struktur das wichtigste Kriterium zur Beurteilung der Lehre dar (Rindermann, 1996, S. 43; Shatz & Best, 1986).

Breite/Bezüge

Beispiele und Darlegungen, die den zu vermittelnden Stoff in verschiedene Kontexte einbetten, erleichtern wie die Strukturierung den Lernprozeß, die Aufnahme, Verarbeitung, Abrufung und aktive Anwendung von Wissen (vgl. Renkl, 1996a). Zu diesen Kontexten zählen Praxisbezüge, Beispiele aus anderen Bereichen und Erläuterungen, welche die Relevanz und Bedeutung der Inhalte aufzeigen.

Verarbeitungstiefe

Lehrziel ist nicht nur die Vermittlung fertigen Wissens, sondern auch die Erreichung höherer kognitiver Ziele (Verständnis, Anwendung, Bewertung). Der Stoff wird als Problem oder Aufgabe, die des eigenen Mit- und Durchdenkens bedarf, dargestellt. Passive Stoffaufnahme erleichtert Vergessensprozesse, während aktive Verarbeitung die Enkodierung und Speicherung auf tieferem Niveau ermöglicht. Zeigt der Dozent z.B. verschiedene Perspektiven auf und nutzt verschiedene Kontexte, fördert er den Verständnisprozeß und die Wissensflexibilisierung (Henninger et al., 1997).

Lehrkompetenz/Rhetorik/Didaktik

Unter Lehrkompetenz wird weniger eine einzelne Unterrichtstechnik, sondern eine allgemeine Fähigkeit zu lehren verstanden. Didaktische Kompetenz äußert sich etwa in guter Vorbereitung, in verständlicher und anregender Sprache (vgl. Döring, 1995, S. 66ff.) oder in der Fähigkeit, Kompliziertes erklären zu können. In angloamerikanischen Verfahren dienen „teaching effectiveness“ und „overall instructor“ oft als übergeordnete Kategorie des auf den Dozenten attribuibaren Varianzanteils der Lehre. In diesem Sinne könnten alle unter dem Dozenten subsumierbaren Bedingungsfaktoren unter Lehrkompetenz gebündelt werden. Didaktische Kompetenz wird nach Marques et al. (1979) über verschiedene Fachbereiche und Urteilergruppen

weitgehend invariant konzeptionalisiert (z.B. Wissensvermittlung, Stimulation von Interesse).

Engagement/Motivierung

Engagement des Dozenten und Motivierung der Teilnehmer gehören wie didaktische Kompetenz zu den wichtigen Attributen guter Hochschullehrer (Feldman, 1976; Klauer, 1985; Marques et al., 1979; Marsh & Roche, 1992; Müller-Wolf, 1977). Desinteresse des Dozenten überträgt sich auf die Teilnehmer einer Veranstaltung, senkt die Aufmerksamkeit und motiviert nicht zu eigener Beschäftigung mit den Themen.

Kooperativität/Klima

Lehre als interpersonale Aktivität erfordert nicht nur fachliche und didaktische Kompetenz, sondern auch soziale Fähigkeiten und Verhaltensweisen wie Freundlichkeit, Kooperativität und Kommunikation (vgl. Moses, 1985).

Interaktion (Förderung, Leitung)

Studentische Beteiligungen und Diskussionen bedürfen neben der Förderung (vgl. Basler et al., 1995; Skala „Motivation zur aktiven Teilnahme“) auch der Leitung: Durch eine freundliche Atmosphäre und offene wie divergente Fragen (Dunkin & Barnes, 1986) werden studentische Beiträge angeregt, die Moderation von Diskussionen verhindert deren thematischen Zerfall. Interaktionsvariablen hängen stärker von Kontextbedingungen der Lehre ab als die Beurteilung der Lehrkompetenz (Frey, 1978).

Betreuung/Feedback

Studierende benötigen neben Rückmeldungen auf Fragen besonders bei der Erstellung von Referaten oder Hausarbeiten Betreuung und Feedback durch den Dozenten. Feedback durch den Dozenten ist auch in Diskussionen zwischen Teilnehmern wichtig (Feldman, 1976). Betreuung sichert die Qualität der Beiträge und stützt die Eigeninitiative. Nach Spies et al. (1996) und Mültrus (1995) stellen Rückmeldungen und die Kontaktdichte zwischen Dozenten und Studierenden wichtige Determinanten der Studienzufriedenheit und der studentischen Bewertung der Studienqualität dar.

Wissenschaftliche Fachkompetenz

Die Fachkompetenz des Dozenten gehört neben der didaktischen, motivationalen und sozialen Kompetenz zu den zentralen Voraussetzungen universitärer Lehre (vgl. Moses, 1985). Diese kann in der Regel nicht durch Studenten beurteilt werden, wenn sie dies könnten, besuchten sie aufgrund der Redundanz die falsche Veranstaltung (Marques et al., 1979). Die Sicherung der fachwissenschaftlichen Qualität stellt eine Aufgabe der Peer-Evaluation dar.

3.3 Bedingungsfaktor Studierende (Kompetenzen, Einstellungen und Handlungen)

Das Handeln des Dozenten stellt nur einen – wenn auch wahrscheinlich den wichtigsten (Müller-Wolf, 1977, S. 115) – Bedingungsfaktor des Erfolges

universitärer Lehre dar. Hinzu müssen auf studentischer Seite spezifische und allgemeine Lernvoraussetzungen sowie veranstaltungsinterne und -externe Handlungen kommen:

Vorwissen

Das Vorwissen der Studenten ist das Ergebnis vorheriger Bildungserfahrungen durch Schule, Universität und Selbststudium. Verfügen die Studierenden über zu wenig, zu viel oder fehlerhaftes Vorwissen, schadet dies dem Lernprozeß und der Lehreffektivität (Renkl, 1996b; Schmidt et al., 1995). Bei zu wenig Wissen müssen erst die Grundlagen erarbeitet werden oder die Themen bleiben bei unzureichender Vor- und Nachbereitung durch die Teilnehmer unverständlich. Bei zu viel Vorbildung tritt Langeweile auf, die – wie Überforderung – zu Störungen oder Besuchsabbruch führen kann. Ist dem Dozenten das Vorwissen bekannt, kann er die Inhalte und Formen der Lehre adaptieren.

Fähigkeiten

Der Lehrprozeß kann nur bei adäquat vorhandenen kognitiven Kompetenzen zu günstigen Lernresultaten führen (Klauer, 1985). Neben diesen Fähigkeiten können auch nichtkognitive Kompetenzen, z.B. Kommunikationskompetenz, Organisationsgeschick oder Frustrationstoleranz, relevant für den Lern- und Studienerfolg sein. Bislang fand jedoch nur die kognitive Leistungsfähigkeit, meist in Form der Abiturnote oder der allgemeinen Intelligenz, als Prädiktorvariable Eingang (Rindermann & Oubaid, 1999), sie korrelierte in bisherigen Studien allerdings nicht mit der Beurteilung der Lehre oder des selbst eingeschätzten Lernerfolges. Prüfungsergebnisse lassen sich aber damit sehr wohl prognostizieren (vgl. Daniel, 1995).

Vorinteresse

Vorinteresse oder Interesse am Thema, d.h. selbstintentionale, positive, wertschätzende Zuwendung, zählt mit Besuchsgrund zu den wichtigsten Bedingungs- und Biasvariablen (Rindermann, 1997a). Unter Biasvariablen sind solche Faktoren zu verstehen, die die Beurteilung der Lehre verzerren können. Sie führen zu einer unzutreffenden Beschreibung der Lehre, die besonders in den Dozentenskalen zu kontrollieren sind. Das Interesse fördert jedoch auch die Lernmotivation und den Lernprozeß (s. Klauer, 1985; Rheinberg, 1996), die Beteiligung und den Fleiß der Studierenden ($r=.34$, Rindermann, 1996, S. 110).

Fleiß/Arbeitshaltung/Arbeitsbelastung

Unter Fleiß oder Arbeitsbelastung fällt die Vor- und Nachbereitung einer Veranstaltung außerhalb der Kurssitzung. Diese sind nicht unabhängig von Rahmenbedingungen wie der Existenz einer Prüfung und der Höhe der Anforderungen (Rindermann, 1996). Mit anderen Skalen ist Fleiß im HILVE unkorreliert (ähnliche Resultate bei Dozenten- und Fremdurteilen). Auch zur Abiturnote und zum Klausurergebnis (individuelle Rohdaten) bestehen keine Zusammenhänge. Fleiß als individuelles Merkmal innerhalb einer Veranstal-

tung (Individualdaten) gibt die persönliche Arbeitshaltung wieder, als Veranstaltungsmittelwert bildet sie eher die von Rahmenbedingungen abhängige Arbeitsbelastung im Kurs ab.

Beteiligung

Ähnlich wie Interaktion und Referate ist die veranstaltungsinterne Beteiligung der Studierenden in Seminaren für den Erfolg der Veranstaltung mitverantwortlich. In Vorlesungen treten mit der Allgemeinbeurteilung Korrelationen um Null auf, in Seminaren wird jedoch $r \approx .50$ erreicht (HILVE-Daten). Die Beteiligung variiert beträchtlich mit Rahmenbedingungen (Veranstaltungstyp, Prüfung) *und* einzelnen didaktischen Skalen (Verarbeitungstiefe; s. Rindermann, 1996). Nach Dunkin und Barnes (1986) ist die studentische Partizipation eine Bedingung für die Förderung höherer kognitiver Prozesse wie kritisches Denken. Durch aktive Einbindung und kooperative Prozesse der Wissensaneignung können Lernprozesse effektiver gestaltet werden, die über den Erwerb assoziativer Inhalte hinausgehen (Seel, 1997).

Referate

Studentische Referate bilden in vielen Seminaren und Tutorien das Fundament der Lehre. Referate dienen als Nachweis aktiver Einarbeitung in ein Themengebiet, vermitteln anderen Studierenden Stoff und stellen Lerngelegenheiten zum Erwerb von Präsentationskompetenz dar (Bromme & Rambow, 1993). Da die Referate den wesentlichen Teil der Inhalte weitergeben und oft die überwiegende Veranstaltungszeit beanspruchen, ist die inhaltliche und didaktische Qualität dieser Beiträge zentral für den Erfolg solcher Veranstaltungen. In Seminaren korreliert die Qualitätsbeurteilung der Referate mit der Allgemeineinschätzung des Kurses zu $r = .60$! Gerade an der Qualität der Referate wird von Studierenden jedoch häufig Kritik geübt, sie werden schlechter beurteilt als das Lehrverhalten des Dozenten (Gold & Mayring, 1997; Rindermann, 1996, S. 30f.). Anhand des HILVE-II lassen sich Referate summativ bewerten (alle vergangenen der Veranstaltung). Zur einzelnen Rückmeldung müßte wie durch Gold und Mayring oder Preiser (1995) ein spezielles, oft einsetzbares Referatfeedbackverfahren angewendet werden.

Störungen

Unruhe und Nebentätigkeiten oder fehlende Pünktlichkeit stören den Veranstaltungsablauf. Störungen sind nicht unabhängig von Rahmenbedingungen (Besuchszahl, Besuchsgrund, Veranstaltungstyp) und von der didaktischen Kompetenz des Dozenten, sie können deshalb als ein Indikator für Defizite in diesen Bereichen gelten.

Fehlzeiten

Auch Fehlzeiten der Studierenden können als Hinweis für Defizite in der Gestaltung der Rahmenbedingungen oder im Lehrverhalten des Dozenten genutzt werden. Allerdings sind diese ebenso von externen Faktoren abhängig, die nicht unter die Kontrolle der Hochschule fallen (Berufstätigkeit, Krankheit).

3.4 Rahmenbedingungen

Thema

Interessantheit und Relevanz des Themas werden in der vorgestellten Konzeption als eine dem Dozenten vorgegebene Rahmenbedingung des Kurses aufgefaßt. Ein von den Studierenden als relevant betrachtetes Thema erleichtert dem Dozenten die Motivierung und fördert die Mitarbeit. Die Beurteilung des Themas korreliert in bisherigen Untersuchungen vor allem mit der Interessantheit der Veranstaltung. Bislang war jedoch die Kausalitätsrichtung nicht feststellbar, da nicht vor der ersten Veranstaltungssitzung anhand eines Seminarplanes das Thema unabhängig vom Kursgeschehen beurteilt wurde (vgl. Rindermann, 1996).

Überschneidungen/Redundanz

Inhaltliche Überschneidungen zwischen Veranstaltungen sind ein Zeichen unzureichender Curriculumsorganisation. Sie können zu geringeren Besuchszahlen oder in Pflichtveranstaltungen zu Nebenbeschäftigungen und Störungen führen.

Anforderungen

Die Höhe der Anforderungen in Form von Stoffschwere, Stoffumfang und Tempo stellt ein zentrales Merkmal zur Beschreibung von Lehrveranstaltungen dar (vgl. Banz & Rodgers, 1985; Bolton et al., 1976; Burdsal & Bardo, 1986). Sie korreliert in HILVE-Studien negativ mit dem selbsteingeschätzten Lernerfolg und positiv mit Fleiß. Allerdings ist ebenso eine negative Beziehung zu Breite/Bezüge erkennbar (Rindermann, 1996, S. 39). In Analysen zur Generalisierbarkeit studentischer Beurteilungen (Rindermann, 1997b) ließ sich beobachten, daß die Beurteilung der Anforderungen anhand des HILVE in bedeutendem Maße mit der Person des Dozenten variiert und somit auch auf sein Lehrverhalten zurückzuführen ist.

Besuchszahl

Steigende Besuchszahl erschwert kommunikative Veranstaltungsformen, z.B. Seminare mit Diskussion und Referaten (McKeachie, 1990). Auch Vorlesungen können bei schlechten räumlichen Bedingungen unter großen Besuchszahlen leiden. Die Beteiligung nimmt in beiden Veranstaltungstypen mit zunehmender Teilnehmerzahl ab (Rindermann, 1996, S. 112; Marsh & Roche, 1992). Andererseits können geringe Besuchszahlen ebenso Folge schlechter Lehre sein. Die Besuchszahl sollte auch zur Ermittlung der Teilnahmequote an der Lehrevaluation durch Relativierung der Fragebogenzahl an der Besuchszahl erhoben werden (Repräsentativitätsmaß).

Prüfung

In Veranstaltungen mit anschließender Prüfung beteiligen sich die Studierenden weniger, sie sind aber fleißiger. Prüfungsvorbereitungskurse zeichnen sich durch hohe Effektivitätsorientierung im Sinne von „Erwerb prüfbaren Wissens pro Zeiteinheit“ aus.

Besuchsgrund

Besuchsgrund gilt neben dem Vorinteresse als wichtigste Biasvariable (Krause, 1998): Kurse, die wegen eines Scheines, einer Prüfung oder schlichtweg wegen Anwesenheitskontrolle zu besuchen sind, werden kritischer beurteilt als Veranstaltungen, die aus Interesse belegt werden (Rindermann, 1997a). Um eine unabhängige Einschätzung zu erhalten, müßte der Besuchsgrund wie die Themenrelevanz vor Veranstaltungsbeginn erhoben werden.

Veranstaltungstyp

Hier werden Vorlesungen, Seminare und praxisorientierte Kurse (Praktika, Übungen) unterschieden. Für Vorlesungen und Seminare, nicht aber für Laborpraktika oder einübende Veranstaltungen (Sport, Sprachübungen), ist das gewählte Modell geeignet. In Vorlesungen wird der fakultative Block Referate frei gelassen, bei Beteiligung, Interaktion und Betreuung treten veridikal niedrige Werte auf. Die Information Veranstaltungstyp ist für die Normierung der Skalen wichtig, die studentische Mitarbeit erheben. Einstellungsänderungen und höhere kognitive Ziele wie Problemlösen sollen nach Dunkin und Barnes (1986) in sozial-interaktiven Lehrformen wie Seminaren eher erreichbar sein als in Vorlesungen (vgl. a. Klauer, 1985; Winteler, 1974).

3.5 Zielgröße Lehrerfolg

Zu den Zielgrößen erfolgreicher Lehre zählen neben veranstaltungsbezogenen Kriterien im engeren Sinne, wie *Interessantheit der Veranstaltungsgestaltung* und *allgemeine Veranstaltungsqualität*, auch Effekte bei den Teilnehmern in Form von quantitativem und qualitativem *Lerngewinn*, *Einstellungsänderungen* und *Kompetenzerwerb* (vgl. Dunkin & Barnes, 1986). Die ersten beiden Effektvariablen auf seiten der Studierenden können wie die Skalen der Bedingungsfaktoren sowohl durch Selbst- als auch durch Fremdbeurteilung erhoben werden. In der hier gewählten Begrifflichkeit wird Lerngewinn als Rating-Variable konzipiert. Als adäquate Maße für den Kompetenzerwerb gelten dagegen Prüfungen und Tests.

Interessantheit der Veranstaltungsgestaltung

Die Interessantheit der Veranstaltungsgestaltung wird als Effektvariable des Dozentenhandelns und im geringeren Maße des studentischen Vorinteresses konzipiert (vgl. empirische Resultate in Rindermann, 1996, 1997b). Manche Forscher heben deren Prozeß-Charakter hervor, da sie die allgemeine Veranstaltungsqualität, den Lernerfolg und Einstellungsänderungen positiv beeinflusst (s. Schiefele, Krapp & Schreyer, 1993).

Allgemeine Veranstaltungsqualität

Wie in den meisten angloamerikanischen Verfahren wird ein „overall course“ aufgenommen (vgl. a. „globale Lehrveranstaltungszufriedenheit“ von Westermann et al., 1996). Diese kann nicht die Ergebnisse in den einzelnen Bedingungsfaktoren ersetzen, vor allem sollten zur Beurteilung des Dozen-

ten primär die Dozentenskalen herangezogen werden und in keinem Falle diffuse Zufriedenheitseinschätzungen (Weiss, 1991). Die allgemeine Veranstaltungsqualität läßt sich jedoch als Kriterium zur Bestimmung der Relevanz einzelner Skalen nutzen.

Lerngewinn (quantitativ/qualitativ)

Der quantitative und qualitative Lerngewinn gilt als wichtigstes Merkmal guter Lehre. Teilnehmer sollten mehr als vorher wissen und sie sollten über ein grundlegendes Verständnis in als wertvoll erachteten Bereichen verfügen (Prosser & Trigwell, 1991). Mit Lerngewinn ist eher die Wissenszunahme im Vergleich zum Veranstaltungsbeginn konzipiert als der interindividuelle Leistungsstand im Vergleich zu anderen Veranstaltungsteilnehmern. Dies läßt sich durch den geringen Zusammenhang zwischen Klausurergebnissen und Lernselbstseinschätzungen empirisch bestätigen ($r=.18$, Rindermann, 1996, S. 120; $r=.13$, Rosemann & Schweer, 1996, S. 178).

Einstellungsänderungen

Gesteigertes Interesse gegenüber den Studieninhalten und eine gewachsene Motivation, sich mit diesen selbständig zu beschäftigen, kennzeichnen hier diese Skala. Gute Lehre weckt Neugier und intrinsisches Interesse (s. Schiefele et al., 1993, S. 123; McKeachie, 1997).

Kompetenzerwerb

Im engeren Sinne bedeutet er der in Prüfungen nachweisbare Lernerfolg. Im weiteren Sinne gehören auch nicht-kognitive Fähigkeiten und Fertigkeiten dazu (z.B. soziale Kompetenzen, Präsentationsmethodik, Lernstrategien). Als Effektvariablen wurden bislang aber nur kognitive Leistungen in Form von Klausurergebnissen herangezogen. Bei Heranziehung *individueller* Daten erwies sich die vorherige Kompetenz (gemessen an der Abiturnote) als wichtigstes Prognosemerkmal für Prüfungserfolg ($r=.48$; vgl. Rindermann, 1996, S. 120). Veranstaltungs- und Dozentenskalen haben im Roh-Datensatz (Vergleich der Studenten untereinander) keine Prognosewirkung. Im Vergleich zwischen Dozenten (gleiche Kurse und Prüfungen, Kursmitteldaten; s. Rindermann, 1997a) korrelieren Dozentenskalen und Klausurergebnisse mit $r=.48$. Bei individuellen Daten kann das Leistungsniveau relativ gut das Klausurergebnis vorhersagen, zwischen Dozenten (bei egalisierten studentischen Stichproben) zeigt sich jedoch die bedeutsame Wirkung des Dozentenverhaltens auf Kursdurchschnittsergebnisse.

Die fünf vorgestellten Lehreffektivitätskriterien, die Ziele universitären Unterrichts wiedergeben, konkurrieren in bestimmten Konstellationen. Beispielsweise ist in „Paukkursen“ mit hohen Arbeitsanforderungen und häufiger Leistungskontrolle kein positiver Einstellungswechsel zu erwarten. Nicht mit jeder Veranstaltungsform sind alle Ziele äquivalent erreichbar, Dozenten sollten je nach Zielpräferenz entsprechende Unterrichtsmethoden einsetzen (vgl. Helmke & Schrader, 1998).

Über diese Skalen hinaus können *fach- und veranstaltungsspezifische Items* die Angemessenheit eines Inventars für verschiedene Lehrziele und Unterrichtsformen sichern (vgl. die Liste freier Items in Rindermann & Amelang, 1994). Erweiterungsmöglichkeiten bestehen z.B. bei den Rahmenbedingungen (Termin, Dauer, Räumlichkeit und Überfüllung), einzelnen Unterrichtsaspekten (Skript, Medien, Hausaufgaben, Arbeit mit Patienten, Prüfungen/Bewertungen) oder der Beurteilung des Fragebogens selbst (Hage, 1996; Marsh, 1987).⁴ Kooperative Arbeitsformen wie Gruppen- und Partnerarbeit sollten hinsichtlich ihrer Gestaltung und Effektivität (z.B. hinsichtlich Prüfungsängstlichkeit und sozialer Integration; Roterding-Steinberg, 1995) bewertet werden. Ein Grunditemsatz deckt die zentralen Unterrichtsprinzipien vieler Veranstaltungen in verschiedenen Studiengängen ab, für spezifische Aspekte ist jedoch eine Adaptation durch hinzufügbare Skalen notwendig.

4. Empirische Überprüfung

Die vorgestellte Konzeption läßt sich mittels verschiedener Methoden überprüfen. Dazu zählen im weiteren Sinne alle statistischen Verfahren, die der Qualitätskontrolle entwickelter Fragebögen dienen (vgl. für das HILVE Rindermann, 1996, 1998): Berechnungen von Urteilerübereinstimmungen zur Einschätzung der Objektivität, von Skalenhomogenitäten und Stabilitäten für die Meßgenauigkeit der Skalen und Items und von Beziehungen zwischen Skalen und externen Kriterien für die Validität. Zusätzlich können Relevanzbefragungen bei Dozenten und Studenten die Praxistauglichkeit sichern. Im engeren Sinne dienen der Konzeptionsprüfung Verfahren, welche die Beziehung zwischen Bedingungs- und Effektvariablen bestimmen, wie schrittweise Regressionen und Pfadanalysen.

Für das Münchner multifaktorielle Modell der Lehrveranstaltungsqualität können zum ersten Mal Ergebnisse empirischer Untersuchungen vorgelegt werden. Stichprobe sind Auszubildende, die Kurse an einer Schule für Logopädie beurteilen, der Veranstaltungscharakter ähnelt Vorlesungen oder dozentenorientierten Seminaren; Referate und Gruppenarbeit finden vergleichsweise selten statt. Bis auf unwesentliche Formulierungsänderungen (statt „Veranstaltung“ „Kurs“ etc.) ist das verwendete Instrument identisch mit dem HILVE-II. Beurteilt wurden 26 Kurse von N=458 Schülern (76% Frauen, 58% Abitur, Alter im Schnitt 32 Jahre). Leistungsdaten in Form von Prüfungen und Klausuren liegen bislang noch nicht vor. Die Besuchszahl variiert kaum (Anwesenheitspflicht, feste Klassengröße von 20 Studenten), sie wurde deshalb nicht als Prädiktorvariable berücksichtigt, das gleiche trifft auf den Veranstaltungstyp und Besuchsgrund zu (alle Kurse Pflichtveranstaltungen und gleicher Veranstaltungstyp).

⁴ Oder Ergänzung durch ephemere Items (z.B. Benachteiligung spezifischer Gruppen, Berücksichtigung bestimmter sozialer oder politischer Themen).

4.1 Prognose des Lehr Erfolgs – Prädiktor-Kriteriums-Beziehungen

Durch schrittweise Regressionsanalysen wird die Prognosekraft von Prädiktorvariablen für Kriteriumsvariablen bestimmt. Prädiktorvariablen sind im vorgestellten Modell Skalen des Dozentenverhaltens, des studentischen Verhaltens und der Rahmenbedingungen. Kriterien sind die einzelnen Lehreffektivitätsskalen (s. Abschnitt 3.5). Prädiktorskalen werden nach ihrem inkrementellen Beitrag in ein statistisches Prognosemodell aufgenommen. Weitere Skalen berücksichtigen Regressionsanalysen nur dann, wenn sie statistisch signifikant Restvarianz aufklären.

Tabelle 1: Prognose des Lehr Erfolgs

Interessantheit der Veranstaltungsgestaltung			
Skala	B	T-Wert	Signifikanz (p)
Struktur/Klarheit	0,33	6,47	.0000
Engagement	0,22	4,22	.0000
Anforderung	0,10	4,05	.0001
Interaktion	0,15	3,79	.0002
Lehrkompetenz	0,20	3,02	.0027
Multiple Korrelation mit signifikanten Prädiktoren: $r=.85$ ($r^2=.73$)			
Allgemeine Veranstaltungsqualität			
Skala	B	T-Wert	Signifikanz (p)
Struktur/Klarheit	0,32	7,50	.0000
Engagement	0,28	6,73	.0000
Verarbeitung	0,16	4,04	.0001
Lehrkompetenz	0,18	3,17	.0016
Multiple Korrelation mit signifikanten Prädiktoren: $r=.89$ ($r^2=.80$)			
Lernen quantitativ			
Skala	B	T-Wert	Signifikanz (p)
Struktur/Klarheit	0,29	6,09	.0000
Engagement	0,32	6,03	.0000
Interaktion	0,14	3,56	.0004
Breite/Bezüge	0,14	3,02	.0026
Multiple Korrelation mit signifikanten Prädiktoren: $r=.82$ ($r^2=.68$)			
Lernen qualitativ			
Skala	B	T-Wert	Signifikanz (p)
Themarelevanz	0,17	6,05	.0000
Engagement	0,28	5,12	.0000
Struktur/Klarheit	0,23	4,16	.0000
Breite/Bezüge	0,22	3,90	.0001
Multiple Korrelation mit signifikanten Prädiktoren: $r=.75$ ($r^2=.56$)			
Interessenförderung			
Skala	B	T-Wert	Signifikanz (p)
Engagement	0,31	5,15	.0000
Verarbeitung	0,23	4,13	.0000
Themarelevanz	0,11	3,83	.0001
Struktur/Klarheit	0,19	3,59	.0004
Fleiß	0,11	3,53	.0004
Vorinteresse	0,09	3,21	.0014
Betreuung	0,11	2,72	.0067
Multiple Korrelation mit signifikanten Prädiktoren: $r=.78$ ($r^2=.61$)			

Anmerkungen: Schrittweise Regressionen, $p<0.01$, $N=458$ (Rohdaten); fehlende Werte durch Mittelwerte ersetzt

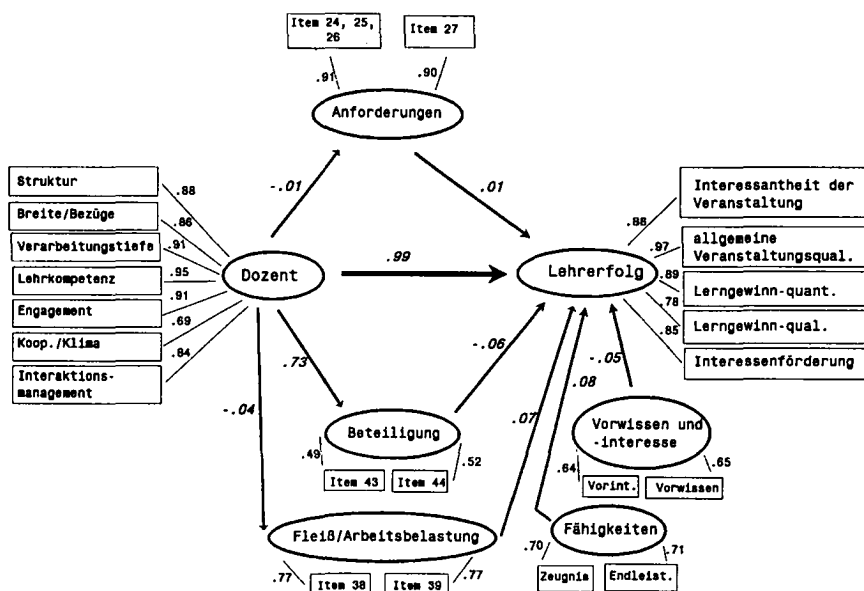
Struktur/Klarheit erweist sich aus Perspektive der Kursteilnehmer als die wichtigste Bedingungsvariable des Lehrerfolges. Bei allen Lehreffektivitätsvariablen wurde sie in die Prognosegleichung aufgenommen. Daneben ist das *Engagement des Dozenten* von großer Relevanz. Studentische Variablen (Vorwissen, Fähigkeiten, Vorinteresse, Fleiß, Beteiligung, Referat, Sörungen und Fehlzeiten) spielen kaum eine Rolle. Nur bei der Interessenförderung sind studentische Variablen von Bedeutung, doch auch hier sind Dozentenengagement und Verarbeitungstiefe relevantere Prognosevariablen. Unter den Rahmenbedingungen sind die Höhe der *Anforderungen* und die *Themarelevanz* bedeutsame Größen, wobei Höhe der Anforderungen und Interessantheit (schwach) positiv miteinander korrelieren ($r=.16$)! Die Beurteilung der Themarelevanz korreliert mit der Selbsteinschätzung, Sinnvolles und Wichtiges zu lernen ($r=.35$). In den Effektivitätsskalen Lernen-qualitativ und Interessenförderung können vom Dozenten nicht kontrollierbare externe oder studentische Bedingungen die Einschätzungen (und den hiermit beschriebenen Unterrichtserfolg) in größerem Ausmaß beeinflussen als bei den anderen drei Effektvariablen. Doch auch hier spielen wie bei Lernen-quantitativ, Allgemeinurteil und Interessantheit die Handlungen des Dozenten eine entscheidende Rolle (Bedingungsvariablen Struktur, Engagement, Verarbeitung).

Die Bedingungsvariablen klären zwischen 56% und 80% der Kriteriumsvarianz auf (im Mittel 68%). Im Vergleich zu Neumann und Finaly (1989), die zwischen 71% und 74% der Lehrerfolgsvarianz durch Prozeßvariablen aufklären konnten, stellt dies eine durchschnittliche Quote dar. Bei Verwendung von Veranstaltungsmitteln wären höhere Werte erwartbar (Fehlerreduktion), bislang sind die HILVE-II-Datensätze hierfür jedoch zu klein.

4.2 Pfadanalysen und Strukturgleichungsmodelle

Strukturgleichungsmodelle dienen der Analyse von Zusammenhangsstrukturen und komplexen Beeinflussungshypothesen auf der Ebene nicht direkt beobachtbarer, latenter Variablen. Kovarianzen lassen sich nur auf Grundlage theoretischer Annahmen als kausale Beziehungen interpretieren. Multidirektionale Modelle wie das vorgestellte Münchner Modell können mittels gängiger Strukturgleichungsmodelle nicht überprüft werden, da jene nur einseitige und nicht reziproke Beeinflussungsrichtungen annehmen (Multidirektionalität wäre nur bei mehreren Beobachtungsintervallen oder bei der Einführung exogener Variablen möglich; Hodapp, 1984, S. 33ff. u. 94ff.). Legt man sich jedoch auf unidirektionale Hauptpfade fest, was – wie jedes Modell und insbesondere jedes statistische Modell – die tatsächliche Systemkomplexität reduziert, sind Modellannahmen mit den bisherigen Daten prüfbar. Hierzu wurden Daten aus der oben angeführten Studie herangezogen ($N=232$ Beobachtungen mit vollständigen Daten in allen Skalen, Rohwerte). Im Gegensatz zu Regressionsanalysen lassen sich nur solche Skalen berücksichtigen, die durch mindestens zwei Items operationalisiert werden. Referate wurden weggelassen, da nur aus wenigen Kursen Einschätzungen vorliegen.

Abbildung 2:
Strukturgleichungsmodell (standardisierte Koeffizienten)



Das *Meßmodell* erreicht bis auf Beteiligung in allen Skalen sehr günstige Werte (Koeffizienten meist bei .80-.90). Die manifesten Variablen sind gute Indikatoren der latenten Variablen. Die Skala Lehrkompetenz ist der beste Indikator aller Dozentenskalen, die Allgemeinschätzung der beste Indikator des Lehrerfolges.

Die meisten *Pfadkoeffizienten* sind von geringer Höhe: Die Beurteilung des Dozenten hat keinen Einfluß auf die Einschätzung der Anforderungshöhe, das gleiche trifft auf Fleiß zu. Auf Beteiligung besteht jedoch seitens des Dozenten (bzw. wie er beurteilt wird) ein großer Effekt. Aber sowohl die Beurteilung der Anforderungen, der Beteiligung, des Fleißes, des Vorwissens und des Vorinteresses sowie der Fähigkeiten bzw. des Leistungsniveaus haben keinen nennenswerten Einfluß auf die Bewertung des Lehrerfolges. Einzig die Einschätzung des Dozenten übt einen großen Effekt auf den Lehrerfolg aus.

Die Gesamtanpassung des Modells beträgt $GFI=.80$ ($RMSEA=0,097$) und ist somit nicht optimal. Dies ist auf Zusammenhänge zwischen den latenten Variablen Beteiligung, Anforderungen und Fleiß zurückzuführen. Modifikationen (Beteiligung als abhängige Variable von Anforderung und Vorwissen/-interesse) erhöhen den GFI geringfügig auf .81 (optimal wäre .90). Aus den beobachtbaren Ergebnissen läßt sich aber ohne Zweifel ablesen, daß die Einschätzung des Dozenten einen weit größeren Einfluß auf die Beurteilung des Lehrerfolges ausübt als Rahmenbedingungsvariablen und studentische Skalen. Die zentrale Determinante des Lehrerfolges in Veranstaltungen ohne be-

deutende studentische Partizipation (z.B. in Form von Referaten) ist aus Teilnehmersicht das Lehrhandeln des Dozenten.

5. Zusammenfassender Ausblick

Das Münchner multifaktorielle Modell der Lehrveranstaltungsqualität stellt eine Weiterentwicklung des Heidelberger Modells dar (Rindermann, 1997a). In der Revision wird nun stärker zwischen den drei Hauptfaktoren Dozent-Studenten-Rahmenbedingungen unterschieden. Auch zählen weitere Skalen wie Interessenförderung, Interaktion oder Vorwissen und die Unterscheidung zwischen Lernen-quantitativ und Lernen-qualitativ zum Grundkanon des revidierten Instruments. Im Gegensatz zu der primär methodisch ausgerichteten angloamerikanischen Lehrevaluations-Forschung liegt dem Münchner Modell ein theoretischer Rahmen zugrunde, welcher die pädagogische Unterrichts-, hochschuldidaktische Lehr- und diagnostisch-methodische Evaluationsforschung zu integrieren versucht. Die auf diesem Modell basierende revidierte Form des Fragebogens wird in Studien in tertiären Ausbildungsinstitutionen (Logopädenausbildung und verschiedenen Fachbereichen an der Universität Frankfurt) eingesetzt. Ein zentrales Resultat bisheriger Analysen mittels schrittweiser Regressionen und Strukturgleichungsverfahren ist, daß sich die Lehreffektivität, gemessen anhand verschiedener Skalen, in weit größerem Ausmaß auf Bedingungsvariablen auf seiten des Dozenten als auf externe Rahmenbedingungen oder studentische Verhaltensweisen und Voraussetzungen zurückführen läßt. Dieses Resultat steht in Übereinstimmung mit Generalisierbarkeitsanalysen anhand des HILVE (Rindermann, 1997b), in denen sich zeigen ließ, daß der Lehrerfolg mit dem Dozenten und seiner Beurteilung über verschiedene Veranstaltungen kovariiert, nicht jedoch mit Veranstaltungsthemen oder Studenten bei verschiedenen Dozenten.

Bevor jedoch abschließende Bewertungen möglich sind, müssen weitere empirische Untersuchungen folgen. Auf methodischer Ebene sind Probleme hinsichtlich des Datenniveaus – Kursmittel sind meßgenauer und aussagekräftiger als Rohdaten (vgl. Rindermann, 1998) – und der Operationalisierung der Bedingungsvariablen zu lösen: Während Dozentenskalen sehr reliabel und homogen meßbar sind, zeigen sich studentische und Rahmenbedingungs-Skalen wesentlich heterogener. Somit wird ihr Einfluß auf den Lehrerfolg in Rechenverfahren schon aus statistischen Gründen reduziert. Ein Fragebogen kann jedoch kaum mehr als 50 Items umfassen, da er im täglichen Einsatz an Hochschulen sonst auf Ablehnung stoßen würde. Ebenso ist es nicht sinnvoll, zur Erhöhung der Meßgenauigkeit sehr ähnliche Items je Skala zu verwenden, diese wären auch aus erhebungsökonomischen Gründen unbrauchbar.

Studentische Daten erlauben nur, Zusammenhänge zwischen Ratings aus studentischer Perspektive zu analysieren. In weiteren Arbeiten sollten deshalb Einschätzungsdaten der Studierenden mit Beobachtungsdaten unterschiedli-

chen Inferenzniveaus (Verhaltensebene, Ratings) von externen Fremdurteilern kombiniert werden. Solche Informationen wie auch experimentelle Variationen, Leistungsdaten und verschiedene instruktionale Settings würden näheren Aufschluß über die Gültigkeit studentischer Urteile und die Zusammenhänge zwischen Bedingungs- und Erfolgsvariablen erlauben. Lehrevaluationsforschung sollte in Zukunft Methoden und Ergebnisse verschiedener Disziplinen (Instruktionspsychologie, Lehr-Lern-Forschung, Hochschuldidaktik, Diagnostik, Wissenschaftssoziologie, Hochschul-, Evaluations- und Interventionsforschung) zusammenführen, um bisherige, sich oft in reiner Bewältigung gesetzlicher Anforderungen erschöpfende Ansätze zu überwinden.

Literatur

- Banz, M. L. & Rodgers, J. L. (1985). Dimensions underlying student ratings of instruction. *American Educational Research Journal*, 22 (2), 267-272.
- Basler, H.-D., Bolm, G., Dickescheid, Th. & Herda, Ch. (1995). Marburger Fragebogen zur Akzeptanz der Lehre. *Diagnostica*, 41(1), 62-79.
- Bolton, B., Bonge, D. & Hinman, S. (1976). Dimensions of students' evaluation of instruction: an empirical synthesis. *Psychological Reports*, 38, 119-123.
- Bromme, R. & Rambow, R. (1993). Verbesserung der mündlichen Präsentation von Referaten. *Das Hochschulwesen*, 6, 289-295.
- Burdsal, Ch. A. & Bardo, J. W. (1986). Measuring student's perceptions of teaching: Dimensions of evaluation. *Educational and Psychological Measurement*, 46, 63-79.
- Cruikshank, D. R. & Kennedy, J. J. (1986). Teacher clarity. *Teaching & Teacher Education*, 2 (1), 43-67.
- Daniel, H.-D. (1995). Der Berufseinstieg von Betriebswirten. *Personal*, 47(10), 492-499.
- Döring, K. W. (1995). *Lehren in der Weiterbildung. Ein Dozentenleitfaden*. Weinheim: Deutscher Studienverlag.
- Dunkin, M. J. & Barnes, J. (1986). Research on teaching in higher education. In M. C. Wittrock (Hg.), *Handbook of research on teaching* (S. 754-777). New York: Macmillan.
- Feldman, K. A. (1976). The superior college teacher from the students' view. *Research in Higher Education*, 5, 243-288.
- Frey, P. W. (1978). A two-dimensional analysis of student ratings of instruction. *Research in Higher Education*, 9, 69-91.
- Gold, A. & Mayring, P. (1997). *Dimensionen studentischer Seminarbewertung*. Frankfurt: Poster auf der 6. Tagung Pädagogische Psychologie.
- Groeben, N. (1986). *Handeln, Tun, Verhalten als Einheiten einer verstehend-erklärenden Psychologie*. Tübingen: Francke.
- Hage, N. el (1996). *Lehrevaluation und studentische Veranstaltungskritik*. Bonn: Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.
- Helmke, A. (1996). Studentische Evaluation der Lehre – Sackgassen und Perspektiven. *Zeitschrift für Pädagogische Psychologie*, 10(3/4), 181-186.
- Helmke, A. & Schrader, F.-W. (1998). Hochschuldidaktik. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 183-187). Weinheim: PVU.

- Henninger, M., Mandl, H. & Balk, M. (1997). Untersuchung eines konstruktivistisch orientierten Trainingsansatzes in der Weiterbildung. *Unterrichtswissenschaft*, 25(4), 365-376.
- Hodapp, V. (1984). *Analyse linearer Kausalmodelle*. Bern: Huber.
- Klauer, K. J. (1985). Framework for a theory of teaching. *Teaching & Teacher Education*, 1(1), 5-17.
- Kramis, J. (1990). Bedeutsamkeit, Effizienz, Lernklima. Grundlegende Gütekriterien für Unterricht und didaktische Prinzipien. *Beiträge zur Lehrerbildung*, 8, 279-296.
- Krause, G. (1998). Besuchsziele – relevante Biasvariablen bei der Lehrevaluation? In G. Krampen & H. Zayer (Hg.), *Psychologiedidaktik und Evaluation* (S. 317-334). Bonn: DPV.
- Krempkow, R. (1998). Ist „gute“ Lehre meßbar? *Das Hochschulwesen*, 46(4), 195-199.
- Land, M. L. (1981). Combined effect of two teacher clarity variables on student achievement. *Journal of Experimental Education*, 50(1), 14-17.
- Leitner, E. (1998). Die Hochschuldidaktik und die Qualität der Hochschullehre. In W. Michl, P. Krupp & Y. Stry (Hg.), *Didaktische Profile der Fachhochschulen* (S. 9-24). Neuwied: Luchterhand.
- Marques, T. E., Lane, D. M. & Dorfman, P. W. (1979). Toward the development of a system for instructional evaluation: Is there consensus regarding what constitutes effective teaching? *Journal of Educational Psychology*, 71(6), 840-849.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.
- Marsh, H. W. & Roche, L. (1992). The use of student evaluations of university teaching in different settings. *Australian Journal of Education*, 36(3), 278-300.
- Marsh, H. W. & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist*, 52(11), 1187-1197.
- Mayr, J. (1997). Evaluieren. In F. Buchberger, H. Eichelberger, K. Klement, J. Mayr, A. Seel & H. Teml (Hg.), *Seminardidaktik* (S. 224-256). Innsbruck: Studienverlag.
- McKeachie, W. J. (1990). Research on college teaching: The historical background. *Journal of Educational Psychology*, 82(2), 189-200.
- McKeachie, W. J. (1997). Student ratings. *American Psychologist*, 52 (11), 1218-1225.
- Moses, I. (1985). High quality teaching in a university: identification and description. *Studies in Higher Education*, 10(3), 301-313.
- Müller-Wolf, H.-M. (1977). *Lehrverhalten an der Hochschule*. München: Verlag Dokumentation.
- Multrus, F. (1995). *Zur Lehr- und Studienqualität. Dimensionen, Skalen und Befunde des Studierendensurveys*. Konstanz: Hefte zur Bildungs- und Hochschulforschung (12).
- Neumann, Y. & Finaly-Neumann, E. (1989). An organizational behavior model of students' evaluation of instruction. *Higher Education*, 18, 227-238.
- Preiser, S. (1995). Feedback nach Referaten. Ein Beitrag zur Verbesserung der Lehre. *Das Hochschulwesen*, 2, 114-116.
- Prosser, M. & Trigwell, K. (1991). Student evaluations of teaching and courses: Student learning approaches and outcomes as criteria of validity. *Contemporary Educational Psychology*, 16, 293-301.
- Reissert, R. (1992). *Evaluation der Lehre – Aktuelle Aktivitäten an deutschen Hochschulen*. Hannover: Hochschul-Informations-System (HIS).

- Renkl, A. (1996a). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, 47, 78-92.
- Renkl, A. (1996b). Vorwissen und Schulleistung. In J. Möller & O. Köller (Hg.), *Emotionen, Kognitionen und Schulleistung* (S. 175-190). Weinheim: Beltz.
- Rheinberg, F. (1996). Von der Lernmotivation zur Lernleistung: Was liegt dazwischen? In J. Möller & O. Köller (Hg.), *Emotionen, Kognitionen und Schulleistung* (S. 23-50). Weinheim: Beltz.
- Rindermann, H. (1996). *Untersuchungen zur Brauchbarkeit studentischer Lehrvaluationen*. Landau: Empirische Pädagogik.
- Rindermann, H. (1997a). Die studentische Beurteilung von Lehrveranstaltungen: Forschungsstand und Implikationen für den Einsatz von Lehrvaluationen. In R. S. Jäger, R. H. Lehmann & G. Trost (Hg.), *Tests und Trends* (S. 12-53). Weinheim: Beltz.
- Rindermann, H. (1997b). Generalisierbarkeit studentischer Veranstaltungsbeurteilungen: Sind Lehrvaluationsresultate generalisierbar auf andere Veranstaltungen eines Dozenten oder auf inhaltsgleiche Parallelveranstaltungen verschiedener Dozenten? *Psychologie in Erziehung und Unterricht*, 44(3), 216-234.
- Rindermann, H. (1998). Übereinstimmung und Divergenz bei der studentischen Beurteilung von Lehrveranstaltungen: Methoden zu ihrer Berechnung und Konsequenzen für die Lehrvaluation. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19(2), 73-92.
- Rindermann, H. & Amelang, M. (1994). *Das Heidelberger Inventar zur Lehrveranstaltungs-Evaluation (HILVE)*. Handanweisung. Heidelberg: Asanger.
- Rindermann, H. & Oubaid, V. (1999). Auswahl von Studienanfängern durch Universitäten – Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 20(3), 172-191.
- Rosemann, B. & Schweer, M. K. W. (1996). Evaluation universitärer Lehre und der Wissenszuwachs bei den Studierenden. *Zeitschrift für Pädagogische Psychologie*, 10(3/4), 175-180.
- Rotering-Steinberg, S. (1995). Kooperative Formen des Lehrens und Lernens in der Erwachsenenbildung. *Unterrichtswissenschaft*, 23(4), 332-346.
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25(2), 120-148.
- Schmidt, H. G., Dolmans, D., Gijselaers, W. H. & Marchais, J. E. (1995). Theory-guided design of a rating scale for course evaluation in problem-based curricula. *Teaching and Learning in Medicine*, 7(2), 82-91.
- Schweer, M. K. W. (1998). Evaluation der Lehre. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 122-125). Weinheim: PVU.
- Seel, A. (1997). Handlungsorientiertes Lernen. In F. Buchberger, H. Eichelberger, K. Klement, J. Mayr, A. Seel & H. Teml (Hg.), *Seminardidaktik* (S. 98-113). Innsbruck: Studienverlag.
- Shatz, M. A. & Best, J. B. (1986). Selection of items for course evaluation by faculty and students. *Psychological Reports*, 58, 239-242.
- Spies, K., Westermann, R., Heise, E. & Schiffler, A. (1996). Diskrepanzen zwischen Bedürfnissen und Angeboten im Studium und ihre Beziehungen zur Studienzufriedenheit. *Empirische Pädagogik*, 10(4), 377-409.
- Webler, W.-D. (1996). Qualitätssicherung in Fachbereichen – ein erster Modellvergleich. *Das Hochschulwesen*, 1, 16-27.

- Weinert, F. E. (1996). Für und Wider die „neuen Lerntheorien“ als Grundlagen pädagogisch-psychologischer Forschung. *Zeitschrift für Pädagogische Psychologie*, 10(1), 1-12.
- Weinert, F. E., Schrader, F.-W. & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, 13, 895-914.
- Weiss, R. (1991). Ziele und Probleme einer Lehrveranstaltungskritik. *Zeitschrift für Hochschuldidaktik*, 15(1-2), 35-42.
- Westermann, R., Heise, E., Spies, K., Hagen, M., Rikert, M., Schiffler, A., Trautwein, U. & Wollburg, S. (1993). *Zufriedenheit mit Studium und Lehrveranstaltungen: Theorieorientierte Konzeptualisierung von Erhebungsverfahren*. Göttingen: Psychologisches Institut, Arbeitsbericht Nr. 1.
- Westermann, R., Heise, E., Spies, K. & Trautwein, U. (1996). Identifikation und Erfassung von Komponenten der Studienzufriedenheit. *Psychologie in Erziehung und Unterricht*, 43, 1-22.
- Wilson, T. C. (1988). Student evaluation-of-teaching forms: A critical perspective. *The Review of Higher Education*, 12(1), 79-95.
- Winteler, A. (1974). *Determinanten der Wirksamkeit akademischer Lehrveranstaltungen*. Meisenheim: Hain.
- Wissenschaftsrat (1996). *Empfehlungen zur Stärkung der Lehre in den Hochschulen durch Evaluation*. Berlin: Bericht des Wissenschaftsrates.

Anschrift des Autors:

Dr. Heiner Rindermann

Institut für Psychologie, Otto-von-Guericke-Universität Magdeburg

Postfach 4120, D-39016 Magdeburg

Telefon 0391/6714832, Fax. 0391/6714815

E-mail: heiner.rindermann@gse-w.uni-magdeburg.de